



# ENTERED

PCT09

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/787,491B

DATE: 02/06/2003 <sup>86</sup>

TIME: 13:49:12

Input Set : A:\pf0600usn\_subseqlist.txt

Output Set: N:\CRF4\02052003\I787491B.raw

#

9

1 <110> APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom;  
 2 CORLEY, Neil C.; GUEGLER, Karl J.;  
 3 GORGONE, Gina A.; PATTERSON, Chandra;  
 4 HILLMAN, Jennifer L.; BAUGHN, Mariah R.;  
 5 LAL, Preeti; AZIMZAI, Yalda;  
 6 YUE, Henry; YANG, Junming  
 8 <120> TITLE OF INVENTION: RNA-ASSOCIATED PROTEINS  
 10 <130> FILE REFERENCE: PF-0600 USN  
 12 <140> CURRENT APPLICATION NUMBER: US 09/787,491B  
 C--> 13 <141> CURRENT FILING DATE: 2002-11-04  
 15 <150> PRIOR APPLICATION NUMBER: PCT/US99/21688  
 16 <151> PRIOR FILING DATE: 1999-09-17  
 18 <150> PRIOR APPLICATION NUMBER: US 60/128,660  
 19 <151> PRIOR FILING DATE: 1999-04-08  
 21 <150> PRIOR APPLICATION NUMBER: US 60/069,391  
 22 <151> PRIOR FILING DATE: 1998-11-04  
 24 <150> PRIOR APPLICATION NUMBER: US 60/183,025  
 25 <151> PRIOR FILING DATE: 1998-09-22  
 27 <150> PRIOR APPLICATION NUMBER: US 60/155,246  
 28 <151> PRIOR FILING DATE: 1998-09-17  
 30 <150> PRIOR APPLICATION NUMBER: US 09/158,720  
 31 <151> PRIOR FILING DATE: 1998-09-17  
 33 <160> NUMBER OF SEQ ID NOS: 38  
 34 <170> SOFTWARE: PERL Program  
 36 <210> SEQ ID NO: 1  
 37 <211> LENGTH: 216  
 38 <212> TYPE: PRT  
 39 <213> ORGANISM: Homo sapiens  
 41 <220> FEATURE:  
 42 <221> NAME/KEY: misc\_feature  
 43 <223> OTHER INFORMATION: Incyte ID No: 399781CD1  
 45 <400> SEQUENCE: 1  
 46 Met Ser Arg Tyr Leu Arg Pro Pro Asn Thr Ser Leu Phe Val Arg  
 47 1 5 10 15  
 48 Asn Val Ala Asp Asp Thr Arg Ser Glu Asp Leu Arg Arg Glu Phe  
 49 20 25 30  
 50 Gly Arg Tyr Gly Pro Ile Val Asp Val Tyr Val Pro Leu Asp Phe  
 51 35 40 45  
 52 Tyr Thr Arg Arg Pro Arg Gly Phe Ala Tyr Val Gln Phe Glu Asp  
 53 50 55 60  
 54 Val Arg Asp Ala Glu Asp Ala Leu His Asn Leu Asp Arg Lys Trp  
 55 65 70 75  
 56 Ile Cys Gly Arg Gln Ile Glu Ile Gln Phe Ala Gln Gly Asp Arg

## RAW SEQUENCE LISTING

DATE: 02/06/2003

PATENT APPLICATION: US/09/787,491B

TIME: 13:49:12

Input Set : A:\pf0600usn\_subseqlist.txt

Output Set: N:\CRF4\02052003\I787491B.raw

```

57          80          85          90
58 Lys Thr Pro Asn Gln Met Lys Ala Lys Glu Gly Arg Asn Val Tyr
59          95          100          105
60 Ser Ser Ser Arg Tyr Asp Asp Tyr Asp Arg Tyr Arg Arg Ser Arg
61          110          115          120
62 Ser Arg Ser Tyr Glu Arg Arg Arg Ser Arg Ser Arg Ser Phe Asp
63          125          130          135
64 Tyr Asn Tyr Arg Arg Ser Tyr Ser Pro Arg Asn Ser Arg Pro Thr
65          140          145          150
66 Gly Arg Pro Arg Arg Arg Glu Ala Ile Pro Thr Met Ile Asp Gln
67          155          160          165
68 Thr Ala Ala Gly Ile Pro Ser Thr Val Leu Leu Thr Thr Leu Gln
69          170          175          180
70 Glu Arg Ser Glu Ser Gly Lys Arg Thr Lys Glu Gly Gln Phe Lys
71          185          190          195
72 Arg Pro Lys Gly Gly Trp Lys Val Leu Gln Tyr Glu Tyr Cys Thr
73          200          205          210
74 Asn Ile Leu Thr Leu Val
75          215
77 <210> SEQ ID NO: 2
78 <211> LENGTH: 1584
79 <212> TYPE: PRT
80 <213> ORGANISM: Homo sapiens
82 <220> FEATURE:
83 <221> NAME/KEY: misc_feature
84 <223> OTHER INFORMATION: Incyte ID No: 1252206CD1
86 <400> SEQUENCE: 2
87 Met Asn Ser Gln Pro Gln Thr Arg Ser Pro Phe Phe Gln Arg Pro
88 1          5          10          15
89 Gln Ile Gln Pro Pro Arg Ala Thr Ile Pro Asn Ser Ser Pro Ser
90          20          25          30
91 Ile Arg Pro Gly Ala Gln Thr Pro Thr Ala Val Tyr Gln Ala Asn
92          35          40          45
93 Gln His Ile Met Met Val Asn His Leu Pro Met Pro Tyr Pro Val
94          50          55          60
95 Pro Gln Gly Pro Gln Tyr Cys Ile Pro Gln Tyr Arg His Ser Gly
96          65          70          75
97 Pro Pro Tyr Val Gly Pro Pro Gln Lys Tyr Pro Val Gln Pro Pro
98          80          85          90
99 Gly Pro Gly Pro Phe Tyr Pro Gly Pro Gly Pro Gly Asp Phe Pro
100          95          100          105
101 Asn Ala Tyr Gly Thr Pro Phe Tyr Pro Ser Gln Pro Val Tyr Gln
102          110          115          120
103 Ser Ala Pro Ile Ile Val Pro Thr Gln Gln Gln Pro Pro Pro Ala
104          125          130          135
105 Lys Arg Glu Lys Lys Thr Ile Arg Ile Arg Asp Pro Asn Gln Gly
106          140          145          150
107 Gly Lys Asp Ile Thr Glu Glu Ile Met Ser Gly Gly Gly Ser Arg
108          155          160          165

```

## RAW SEQUENCE LISTING

DATE: 02/06/2003

PATENT APPLICATION: US/09/787,491B

TIME: 13:49:12

Input Set : A:\pf0600usn\_subseqlist.txt

Output Set: N:\CRF4\02052003\I787491B.raw

109	Asn	Pro	Thr	Pro	Pro	Ile	Gly	Arg	Pro	Thr	Ser	Thr	Pro	Thr	Pro
110					170					175					180
111	Pro	Gln	Leu	Pro	Ser	Gln	Val	Pro	Glu	His	Ser	Pro	Val	Val	Tyr
112					185					190					195
113	Gly	Thr	Val	Glu	Ser	Ala	His	Leu	Ala	Ser	Thr	Pro	Val	Thr	
114					200					205					210
115	Ala	Ala	Ser	Asp	Gln	Lys	Gln	Glu	Glu	Lys	Pro	Lys	Pro	Asp	Pro
116					215					220					225
117	Val	Leu	Lys	Ser	Pro	Ser	Pro	Val	Leu	Arg	Leu	Val	Leu	Ser	Gly
118					230					235					240
119	Glu	Lys	Lys	Glu	Gln	Glu	Gly	Gln	Thr	Ser	Glu	Thr	Thr	Ala	Ile
120					245					250					255
121	Val	Ser	Ile	Ala	Glu	Leu	Pro	Leu	Pro	Pro	Ser	Pro	Thr	Thr	Val
122					260					265					270
123	Ser	Ser	Val	Ala	Arg	Ser	Thr	Ile	Ala	Ala	Pro	Thr	Ser	Ser	Ala
124					275					280					285
125	Leu	Ser	Ser	Gln	Pro	Ile	Phe	Thr	Thr	Ala	Ile	Asp	Asp	Arg	Cys
126					290					295					300
127	Glu	Leu	Ser	Ser	Pro	Arg	Glu	Asp	Thr	Ile	Pro	Ile	Pro	Ser	Leu
128					305					310					315
129	Thr	Ser	Cys	Thr	Glu	Thr	Ser	Asp	Pro	Leu	Pro	Thr	Asn	Glu	Asn
130					320					325					330
131	Asp	Asp	Asp	Ile	Cys	Lys	Lys	Pro	Cys	Ser	Val	Ala	Pro	Asn	Asp
132					335					340					345
133	Ile	Pro	Leu	Val	Ser	Ser	Thr	Asn	Leu	Ile	Asn	Glu	Ile	Asn	Gly
134					350					355					360
135	Val	Ser	Glu	Lys	Leu	Ser	Ala	Thr	Glu	Ser	Ile	Val	Glu	Ile	Val
136					365					370					375
137	Lys	Gln	Glu	Val	Leu	Pro	Leu	Thr	Leu	Glu	Leu	Glu	Ile	Leu	Glu
138					380					385					390
139	Asn	Pro	Pro	Glu	Glu	Met	Lys	Leu	Glu	Cys	Ile	Pro	Ala	Pro	Ile
140					395					400					405
141	Thr	Pro	Ser	Thr	Val	Pro	Ser	Phe	Pro	Pro	Thr	Pro	Pro	Thr	Pro
142					410					415					420
143	Pro	Ala	Ser	Pro	Pro	His	Thr	Pro	Val	Ile	Val	Pro	Ala	Ala	Ala
144					425					430					435
145	Thr	Thr	Val	Ser	Ser	Pro	Ser	Ala	Ala	Ile	Thr	Val	Gln	Arg	Val
146					440					445					450
147	Leu	Glu	Glu	Asp	Glu	Ser	Ile	Arg	Thr	Cys	Leu	Ser	Glu	Asp	Ala
148					455					460					465
149	Lys	Glu	Ile	Gln	Asn	Lys	Ile	Glu	Val	Glu	Ala	Asp	Gly	Gln	Thr
150					470					475					480
151	Glu	Glu	Ile	Leu	Asp	Ser	Gln	Asn	Leu	Asn	Ser	Arg	Arg	Ser	Pro
152					485					490					495
153	Val	Pro	Ala	Gln	Ile	Ala	Ile	Thr	Val	Pro	Lys	Thr	Trp	Lys	Lys
154					500					505					510
155	Pro	Lys	Asp	Arg	Thr	Arg	Thr	Thr	Glu	Glu	Met	Leu	Glu	Ala	Glu
156					515					520					525
157	Leu	Glu	Leu	Lys	Ala	Glu	Glu	Glu	Leu	Ser	Ile	Asp	Lys	Val	Leu

## RAW SEQUENCE LISTING

DATE: 02/06/2003

PATENT APPLICATION: US/09/787,491B

TIME: 13:49:12

Input Set : A:\pf0600usn\_subseqlist.txt

Output Set: N:\CRF4\02052003\I787491B.raw

158		530		535		540
159	Glu Ser Glu Gln Asp Lys Met Ser Gln Gly Phe His Pro Glu Arg					
160		545		550		555
161	Asp Pro Ser Asp Leu Lys Lys Val Lys Ala Val Glu Glu Asn Gly					
162		560		565		570
163	Glu Glu Ala Glu Pro Val Arg Asn Gly Ala Glu Ser Val Ser Glu					
164		575		580		585
165	Gly Glu Gly Ile Asp Ala Asn Ser Gly Ser Thr Asp Ser Ser Gly					
166		590		595		600
167	Asp Gly Val Thr Phe Pro Phe Lys Pro Glu Ser Trp Lys Pro Thr					
168		605		610		615
169	Asp Thr Glu Gly Lys Lys Gln Tyr Asp Arg Glu Phe Leu Leu Asp					
170		620		625		630
171	Phe Gln Phe Met Pro Ala Cys Ile Gln Lys Pro Glu Gly Leu Pro					
172		635		640		645
173	Pro Ile Ser Asp Val Val Leu Asp Lys Ile Asn Gln Pro Lys Leu					
174		650		655		660
175	Pro Met Arg Thr Leu Asp Pro Arg Ile Leu Pro Arg Gly Pro Asp					
176		665		670		675
177	Phe Thr Pro Ala Phe Ala Asp Phe Gly Arg Gln Thr Pro Gly Gly					
178		680		685		690
179	Arg Gly Val Pro Leu Leu Asn Val Gly Ser Arg Arg Ser Gln Pro					
180		695		700		705
181	Gly Gln Arg Arg Glu Pro Arg Lys Ile Ile Thr Val Ser Val Lys					
182		710		715		720
183	Glu Asp Val His Leu Lys Lys Ala Glu Asn Ala Trp Lys Pro Ser					
184		725		730		735
185	Gln Lys Arg Asp Ser Gln Ala Asp Asp Pro Glu Asn Ile Lys Thr					
186		740		745		750
187	Gln Glu Leu Phe Arg Lys Val Arg Ser Ile Leu Asn Lys Leu Thr					
188		755		760		765
189	Pro Gln Met Phe Asn Gln Leu Met Lys Gln Val Ser Gly Leu Thr					
190		770		775		780
191	Val Asp Thr Glu Glu Arg Leu Lys Gly Val Ile Asp Leu Val Phe					
192		785		790		795
193	Glu Lys Ala Ile Asp Glu Pro Ser Phe Ser Val Ala Tyr Ala Asn					
194		800		805		810
195	Met Cys Arg Cys Leu Val Thr Leu Lys Val Pro Met Ala Asp Lys					
196		815		820		825
197	Pro Gly Asn Thr Val Asn Phe Arg Lys Leu Leu Leu Asn Arg Cys					
198		830		835		840
199	Gln Lys Glu Phe Glu Lys Asp Lys Ala Asp Asp Asp Val Phe Glu					
200		845		850		855
201	Lys Lys Gln Lys Glu Leu Glu Ala Ala Ser Ala Pro Glu Glu Arg					
202		860		865		870
203	Thr Arg Leu His Asp Glu Leu Glu Glu Ala Lys Asp Lys Ala Arg					
204		875		880		885
205	Arg Arg Ser Ile Gly Asn Ile Lys Phe Ile Gly Glu Leu Phe Lys					
206		890		895		900

## RAW SEQUENCE LISTING

DATE: 02/06/2003

PATENT APPLICATION: US/09/787,491B

TIME: 13:49:12

Input Set : A:\pf0600usn\_subseqlist.txt

Output Set: N:\CRF4\02052003\I787491B.raw

207	Leu	Lys	Met	Leu	Thr	Glu	Ala	Ile	Met	His	Asp	Cys	Val	Val	Lys
208					905					910					915
209	Leu	Leu	Lys	Asn	His	Asp	Glu	Glu	Ser	Leu	Glu	Cys	Leu	Cys	Arg
210					920					925					930
211	Leu	Leu	Thr	Thr	Ile	Gly	Lys	Asp	Leu	Asp	Phe	Glu	Lys	Ala	Lys
212					935					940					945
213	Pro	Arg	Met	Asp	Gln	Tyr	Phe	Asn	Gln	Met	Glu	Lys	Ile	Val	Lys
214					950					955					960
215	Glu	Lys	Lys	Thr	Ser	Ser	Arg	Ile	Arg	Phe	Met	Leu	Gln	Asp	Val
216					965					970					975
217	Ile	Asp	Leu	Arg	Leu	Cys	Asn	Trp	Val	Ser	Arg	Arg	Ala	Asp	Gln
218					980					985					990
219	Gly	Pro	Lys	Thr	Ile	Glu	Gln	Ile	His	Lys	Glu	Ala	Lys	Ile	Glu
220					995					1000					1005
221	Glu	Gln	Glu	Glu	Gln	Arg	Lys	Val	Gln	Gln	Leu	Met	Thr	Lys	Glu
222					1010					1015					1020
223	Lys	Arg	Arg	Pro	Gly	Val	Gln	Arg	Val	Asp	Glu	Gly	Gly	Trp	Asn
224					1025					1030					1035
225	Thr	Val	Gln	Gly	Ala	Lys	Asn	Ser	Arg	Val	Leu	Asp	Pro	Ser	Lys
226					1040					1045					1050
227	Phe	Leu	Lys	Ile	Thr	Lys	Pro	Thr	Ile	Asp	Glu	Lys	Ile	Gln	Leu
228					1055					1060					1065
229	Val	Pro	Lys	Ala	Gln	Leu	Gly	Ser	Trp	Gly	Lys	Gly	Ser	Ser	Gly
230					1070					1075					1080
231	Gly	Ala	Lys	Ala	Ser	Glu	Thr	Asp	Ala	Leu	Arg	Ser	Ser	Ala	Ser
232					1085					1090					1095
233	Ser	Leu	Asn	Arg	Phe	Ser	Ala	Leu	Gln	Pro	Pro	Ala	Pro	Ser	Gly
234					1100					1105					1110
235	Ser	Thr	Pro	Ser	Thr	Pro	Val	Glu	Phe	Asp	Ser	Arg	Arg	Thr	Leu
236					1115					1120					1125
237	Thr	Ser	Arg	Gly	Ser	Met	Gly	Arg	Glu	Lys	Asn	Asp	Lys	Pro	Leu
238					1130					1135					1140
239	Pro	Ser	Ala	Thr	Ala	Arg	Pro	Asn	Thr	Phe	Met	Arg	Gly	Gly	Ser
240					1145					1150					1155
241	Ser	Lys	Asp	Leu	Leu	Asp	Asn	Gln	Ser	Gln	Glu	Glu	Gln	Arg	Arg
242					1160					1165					1170
243	Glu	Met	Leu	Glu	Thr	Val	Lys	Gln	Leu	Thr	Gly	Gly	Val	Asp	Val
244					1175					1180					1185
245	Glu	Arg	Asn	Ser	Thr	Glu	Ala	Glu	Arg	Asn	Lys	Thr	Arg	Glu	Ser
246					1190					1195					1200
247	Ala	Lys	Pro	Glu	Ile	Ser	Ala	Met	Ser	Ala	His	Asp	Lys	Ala	Ala
248					1205					1210					1215
249	Leu	Ser	Glu	Glu	Glu	Leu	Glu	Arg	Lys	Ser	Lys	Ser	Ile	Ile	Asp
250					1220					1225					1230
251	Glu	Phe	Leu	His	Ile	Asn	Asp	Phe	Lys	Glu	Ala	Met	Gln	Cys	Val
252					1235					1240					1245
253	Glu	Glu	Leu	Asn	Ala	Gln	Gly	Leu	Leu	His	Val	Phe	Val	Arg	Val
254					1250					1255					1260
255	Gly	Val	Glu	Ser	Thr	Leu	Glu	Arg	Ser	Gln	Ile	Thr	Arg	Asp	His

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/787,491B

DATE: 02/06/2003  
TIME: 13:49:13

Input Set : A:\pf0600usn\_subseqlist.txt  
Output Set: N:\CRF4\02052003\I787491B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:26; N Pos. 1437

**VERIFICATION SUMMARY**

DATE: 02/06/2003

PATENT APPLICATION: US/09/787,491B

TIME: 13:49:13

Input Set : A:\pf0600usn\_subseqlist.txt

Output Set: N:\CRF4\02052003\I787491B.raw

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:1659 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26 after pos.:1380